

ABSTRACT

A unified clustering tree (500) generates phoneme clusters based on an input sequence of phonemes. The number of possible clusters is significantly less than the number of possible combinations of input phonemes. Nodes (510, 511) in the unified clustering tree are arranged into levels such that the clustering tree generates clusters for multiple speech recognition models. Models that correspond to higher levels in the unified clustering tree are coarse models relative to more fine-grain models at lower levels of the clustering tree.